

Efficient method for computing lower bounds on the p -radius of switched linear systems

Masaki Ogura^{a,*}, Victor M. Preciado^a, Raphaël M. Jungers^{b,c}

^a*Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA 19104, USA*

^b*ICTEAM Institute, Université catholique de Louvain, 4 avenue Georges Lemaître, B-1348 Louvain-la-Neuve, Belgium*

^c*F.R.S.-FNRS Research Associate.*

Abstract

This paper proposes lower bounds on a quantity called L^p -norm joint spectral radius, or in short, p -radius, of a finite set of matrices. Despite its wide range of applications to, for example, stability analysis of switched linear systems and the equilibrium analysis of switched linear economical models, algorithms for computing the p -radius are only available in a very limited number of particular cases. The proposed lower bounds are given as the spectral radius of an average of the given matrices weighted via Kronecker products and do not place any requirements on the set of matrices. We show that the proposed lower bounds theoretically extend and also can practically improve the existing lower bounds. A Markovian extension of the proposed lower bounds is also presented.

Keywords: p -radius, switched linear systems, mean stability, Markov processes

1. Introduction

The L^p -norm joint spectral radius [1, 2] (often called p -radius) of an indexed family of $n \times n$ real matrices $\mathcal{M} = \{A_1, \dots, A_N\}$ is defined by, for

*Corresponding author

Email addresses: `ogura@seas.upenn.edu` (Masaki Ogura),
`preciado@seas.upenn.edu` (Victor M. Preciado), `raphael.jungers@uclouvain.be`
 (Raphaël M. Jungers)

$p \geq 1$,

$$\rho_p(\mathcal{M}) = \lim_{k \rightarrow \infty} \left(\frac{\sum_{i_1, \dots, i_k=1}^N \|A_{i_k} \cdots A_{i_1}\|^p}{N^k} \right)^{1/(pk)},$$

where $\|\cdot\|$ denotes the maximum singular value of a matrix. Introduced by Jia [3] and Wang [4] independently for $p = 1$, the p -radius now plays an important role in various fields of applied mathematics. A classical application of the p -radius is in the characterization of the regularity of wavelet functions in L^p spaces [4, 1]. The mean stability of a class of switched linear systems is determined by the value of the p -radius [5, 6]. The so-called indeterminacy of a switching linear economic model can be checked through the p -radius [7].

However, being defined as a limit of a sequence on matrix products, the p -radius is known to be difficult to compute. There is no formula available for its computation except under the conditions that p is an even integer or that the given set of matrices leaves a common proper cone invariant [2, 5]. Although for the latter case there exists a converging approximation method [8] that does not require p to be an integer, in a general case, even approximating the p -radius is an NP-hard problem [8]. As for bounds on the p -radius, the sequence defining the p -radius is decreasing and therefore gives upper bounds, though their computation requires exponentially growing costs. Finally, the lower bounds on the p -radius in the literature [9, 7] are often not very accurate.

In this paper we propose novel lower bounds on the p -radius for integer values of p with no assumptions on the given set of matrices on the contrary to [2, 8]. The lower bounds are given as the spectral radius of a weighted average of the given matrices and the weights are realized by Kronecker product of matrices. We will show that, with appropriately chosen weighting matrices, the proposed bounds extend and also can improve the lower bounds in the literature [9, 7]. The obtained results are furthermore generalized to the Markovian setting. This in particular enables us to use the p -radius to study the stability of so-called Markov jump linear systems [10], which are switched linear systems whose parameter changes by following a time-homogeneous Markov chain. The generalization is based on a stochastic counterpart of the so-called Ω -lift of matrices [11].

The organization of this paper is as follows. In Section 2, we derive a novel lower bound for the p -radius. In Section 3, we provide a Markovian extension. The notations used in this paper are standard. The identity matrix is denoted by I . The spectral radius of a square matrix is denoted

by $\rho(\cdot)$. The Kronecker product (see, e.g., [12]) of matrices A and B is denoted by $A \otimes B$. \mathcal{M} denotes an indexed family $\{A_1, \dots, A_N\}$ of $n \times n$ real matrices. Finally, for ease of reference, we list some important properties of the p -radius in the following proposition:

Proposition 1. *Let $p \geq 1$ be a positive integer.*

1. *The sequence $\{h_k(\mathcal{M})\}_{k=1}^\infty$ defined by*

$$h_k(\mathcal{M}) = \left(\frac{\sum_{i_1, \dots, i_k=1}^N \|A_{i_k} \cdots A_{i_1}\|^p}{N^k} \right)^{1/(pk)}$$

is decreasing [8].

2. *If either p is even or \mathcal{M} leaves a proper cone invariant, then $\rho_p(\mathcal{M}) = \rho(N^{-1} \sum_{i=1}^N A_i^{\otimes p})^{1/p}$, where $A^{\otimes p}$ denotes the Kronecker product of p copies of A (see [2]).*

3. *It holds that $\rho_p(A_1, \dots, A_N) = (\rho_1(A_1^{\otimes p}, \dots, A_N^{\otimes p}))^{1/p}$ (see [8]).*

Remark 2. Throughout the paper, we often omit the dependence of quantities on the underlying family \mathcal{M} of matrices when it is clear from the context. Precisely speaking, for a function f defined on $(\mathbb{R}^{n \times n})^N$, we may simply write $f(\mathcal{M})$ as f . Also, abusing notation, we sometimes write $f(\mathcal{M})$ as $f(A_1, \dots, A_N)$ when $\mathcal{M} = \{A_1, \dots, A_N\}$.

2. Novel lower bounds

In this section, we present novel lower bounds on the p -radius. We also prove that these new bounds outperform existing lower bounds. Notice that, according to the third claim of Proposition 1, any result on the 1-radius ($p = 1$) is directly applicable to the general p -radius; thus, we shall focus on the particular case $p = 1$ for the rest of the paper. In order to state our main results, we need to recall the definition of the joint spectral radius [13]:

$$\rho_\infty(\mathcal{M}) = \limsup_{k \rightarrow \infty} \max_{i_1, \dots, i_k \in \{1, \dots, N\}} \|A_{i_k} \cdots A_{i_1}\|^{1/k}.$$

Our first result is stated in the next theorem:

Theorem 3. For $\mathcal{W} = \{W_1, \dots, W_N\} \subset \mathbb{R}^{m \times m}$, let

$$\lambda_{\mathcal{W}}(\mathcal{M}) = \rho \left(\frac{1}{N} \sum_{i=1}^N W_i \otimes A_i \right).$$

If $\rho_{\infty}(\mathcal{W}) = 1$, then $\lambda_{\mathcal{W}}(\mathcal{M}) \leq \rho_1(\mathcal{M})$.

Proof. Using Gelfand's formula, we obtain

$$\lambda_{\mathcal{W}}(\mathcal{M}) = \lim_{k \rightarrow \infty} \left\| \left(\frac{1}{N} \sum_{i=1}^N W_i \otimes A_i \right)^k \right\|^{1/k}. \quad (1)$$

Using general identities [12] about Kronecker products:

$$(A \otimes C)(B \otimes D) = (AB) \otimes (CD), \quad (2)$$

$$\|A \otimes B\| = \|A\| \|B\|, \quad (3)$$

we can evaluate the norm in (1) as

$$\begin{aligned} \left\| \left(\frac{1}{N} \sum_{i=1}^N W_i \otimes A_i \right)^k \right\| &= \left\| \sum_{i_1, \dots, i_k=1}^N \frac{(W_{i_k} \otimes A_{i_k}) \cdots (W_{i_1} \otimes A_{i_1})}{N^k} \right\| \\ &\leq \sum_{i_1, \dots, i_k=1}^N \frac{\|(W_{i_k} \otimes A_{i_k}) \cdots (W_{i_1} \otimes A_{i_1})\|}{N^k} \\ &= \sum_{i_1, \dots, i_k=1}^N \frac{\|W_{i_k} \cdots W_{i_1}\| \|A_{i_k} \cdots A_{i_1}\|}{N^k} \\ &\leq \left(\max_{i_1, \dots, i_k} \|W_{i_k} \cdots W_{i_1}\| \right) \frac{\sum_{i_1, \dots, i_k} \|A_{i_k} \cdots A_{i_1}\|}{N^k}. \end{aligned}$$

Taking the power of $1/k$ of the last expression, and substituting in (1), we obtain

$$\lambda_{\mathcal{W}}(\mathcal{M}) \leq \rho_{\infty}(\mathcal{W}) \rho_1(\mathcal{M}) = \rho_1(\mathcal{M}), \quad (4)$$

as desired. \square

Remark 4. From (4), we can see that the equality constraint $\rho_{\infty}(\mathcal{W}) = 1$ in Theorem 3 can be relaxed to the inequality constraint $\rho_{\infty}(\mathcal{W}) \leq 1$. Although

the largest lower bound is clearly attained when the joint spectral radius of \mathcal{W} is equal to one, it is convenient in practice to relax the equality constraint because checking that $\rho_\infty(\mathcal{W}) = 1$ is NP-hard [13]. In contrast, checking the inequality $\rho_\infty(\mathcal{W}) \leq 1$ can be efficiently done (using, for example, the JSR software toolbox [14]).

In what follows, we compare the proposed lower bounds with the most relevant existing bounds found in the literature. The first lower bound is the one implicitly obtained by Zhou [9]; if we let

$$\ell_Z = \frac{\rho(\sum_{i=1}^N A_i \otimes A_i)}{N\rho_\infty},$$

then $\rho_1 \geq \ell_Z$. This inequality can be derived from the following bound on the joint spectral radius: $\rho_\infty \geq (\rho_{p+q}/\rho_p)^{p/q} \rho_{p+q}$ whenever $p, q \geq 1$ [9, p. 48]. Letting $p = q = 1$ gives $\rho_1 \geq \rho_2^2/\rho_\infty$. Then, applying assertion 2 of Proposition 1 to this inequality proves $\rho_1 \geq \ell_Z$. The second lower bound was introduced in [7]: for $w_1, \dots, w_N \in [-1, 1]$ it holds that

$$\rho_1 \geq \rho \left(\frac{1}{N} \sum_{i=1}^N w_i A_i \right).$$

In the following theorem, we show that the lower bound herein proposed extends the ones mentioned above:

Theorem 5. *For each $m \geq 1$ define*

$$\ell_m(\mathcal{M}) = \sup\{\lambda_{\mathcal{W}} : \mathcal{W} \in (\mathbb{R}^{m \times m})^N, \rho_\infty(\mathcal{W}) = 1\}. \quad (5)$$

The following statements are true:

1. *If $m \geq n$, then $\ell_m \geq \ell_Z$.*
2. *Define*

$$\ell_{[-1,1]} = \max_{w_1, \dots, w_N \in [-1,1]} \rho \left(\frac{1}{N} \sum_{i=1}^N w_i A_i \right). \quad (6)$$

Then $\ell_{[-1,1]} \leq \ell_m$ for every $m \geq 1$.

Proof. We begin by showing that

$$[m < m'] \Rightarrow [\ell_m \leq \ell_{m'}]. \quad (7)$$

Let us take an arbitrary $\mathcal{W} \subset (\mathbb{R}^{m \times m})^N$ such that $\rho_\infty(\mathcal{W}) = 1$. For each i we define the block diagonal matrix $V_i = \text{diag}(W_i, O_{m'-m}) \in \mathbb{R}^{m' \times m'}$ and let $\mathcal{V} = \{V_1, \dots, V_N\}$. Then, it holds that $\lambda_{\mathcal{W}} = \lambda_{\mathcal{V}}$. Since $\rho_\infty(\mathcal{V}) = 1$, this implies that $\lambda_{\mathcal{W}} \leq \ell_{m'}$. Finally, taking the supremum with respect to \mathcal{W} in the last inequality proves (7).

Let us prove the theorem. It is clear that the normalized family of matrices $\mathcal{M}_0 = \{A_1/\rho_\infty, \dots, A_N/\rho_\infty\} \subset \mathbb{R}^{n \times n}$ has a joint spectral radius equal to one. Therefore, from Theorem 3 it follows that $\ell_n \geq \lambda_{\mathcal{M}_0} = \rho(N^{-1} \sum_{i=1}^N A_i \otimes A_i)/\rho_\infty = \ell_Z$. This inequality and (7) prove the first claim in the theorem. To prove the second claim, we take an arbitrary family $\{w_1, \dots, w_N\} \subset [-1, 1]$. Since this family has a joint spectral radius less than or equal to one, we obtain $\rho(N^{-1} \sum_{i=1}^N w_i A_i) = \lambda_{\{w_1, \dots, w_N\}} \leq \ell_1$ by Theorem 3 and Remark 4. Therefore, $\ell_{[-1, 1]} \leq \ell_1$ and hence (7) proves the second claim. We finally remark that the maximum in (6) exists because the set $[-1, 1]^N$ is compact and the function $\rho(\cdot)$ is continuous. \square

In what follows, we illustrate our results with some examples.

Example 6. Let $N \geq 1$ be arbitrary. Consider the matrix family $\mathcal{M} = \{A_1, \dots, A_{N+1}\}$, where $A_1 = NI$ and

$$A_2 = \dots = A_{N+1} = R = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \quad (8)$$

Since $\|R^k\| = 1$ for every $k \geq 0$, we see that ρ_1 equals the 1-radius of the family of scalars $\{N, 1, \dots, 1\}$, where element 1 has multiplicity N . The 1-radius of this family equals

$$\rho\left(\frac{N+1+\dots+1}{N+1}\right) = \frac{2N}{N+1}$$

by assertion 2 of Proposition 1. Therefore $\rho_1(\mathcal{M}) = (2N)/(N+1)$. This value is attained by the proposed lower bound ℓ_2 because

$$\begin{aligned} \lambda_{\{I, R, \dots, R\}} &= \rho\left(\frac{1}{N+1} (I \otimes (NI) + NR \otimes R)\right) \\ &= \frac{N}{N+1} \rho(I \otimes I + R \otimes R) \\ &= \frac{2N}{N+1}. \end{aligned}$$

On the other hand, a straightforward computation shows that $\ell_Z = 1$.

Example 7. Let $\mathcal{M} = \{I, R, R^2, R^3\}$, where R is given in (8). Clearly, $\rho_1 = 1$ because $\|R^k\| = 1$ for every $k \geq 0$. The proposed lower bound attains this exact value of the 1-radius as $\ell_2 = 1$ because $\lambda_{I, R, R^2, R^3} = 1$. On the other hand, we can show that $\ell_{[-1,1]} \leq \sqrt{2}/2$ as follows. Let $w_i \in [-1, 1]$ ($i = 1, \dots, 4$) be arbitrary. Then we have

$$S = \frac{w_1 I + w_2 R + w_3 R^2 + w_4 R^3}{4} = \frac{1}{4} \begin{bmatrix} u & -v \\ v & u \end{bmatrix},$$

where $u = w_1 - w_3$ and $v = w_2 - w_4$. Since $|u| \leq 2$ and $|v| \leq 2$, we can see $\rho(S) \leq \sqrt{2}/2$ and therefore $\ell_{[-1,1]} \leq \sqrt{2}/2$.

Although ℓ_m can largely improve other lower bounds in the literature, it is not easy to compute. The first reason is the non-convexity of the function ρ (see, e.g., [15]). The other reason is that the set $\{\mathcal{W} \in (\mathbb{R}^{m \times m})^N : \rho_\infty(\mathcal{W}) = 1\}$ does not admit an appropriate parametrization due to the NP-hardness [13] of computing ρ_∞ . For these reasons, we here propose using the set of matrix weights from the following set:

$$\mathfrak{D}_m = \left\{ \{D_i L_i\}_{i=1}^N : D_i \text{ is diagonal, } \|D_i\| \leq 1, \text{ and } L_i \text{ is orthogonal} \right\}.$$

Since the matrices in this set have a joint spectral radius less than or equal to one, the best lower bound $\ell_{\mathfrak{D}_m} = \max\{\lambda_{\mathcal{W}} : \mathcal{W} \in \mathfrak{D}_m\}$ achieved by using such matrix weights provides a lower bound on the 1-radius by Theorem 3 and Remark 4. Moreover, since an orthogonal matrix L admits [16] the parametrization $L = D(I - S)(I + S)^{-1}$ where S is a skew-symmetric matrix and D is a diagonal matrix whose diagonals are either $+1$ or -1 , we can maximize $\ell_{\mathfrak{D}_m}$ using, for example, a stochastic gradient descent algorithm [17]. Notice that the result of this algorithm will be a local maximum.

In the following example, we illustrate the effectiveness of the weights from the set \mathfrak{D}_m by studying the stability of a switched linear system.

Example 8. Consider the switched linear system

$$X(k+1) = A_{\sigma(k+1)} X(k), \quad X(0) = I \quad (9)$$

where $\{\sigma(k)\}_{k=1}^\infty$ are random variables independently and uniformly distributed on $\{1, \dots, N\}$. The system is said to be *p*th mean stable [18] if there exist $C > 0$ and $\gamma \in [0, 1)$ such that

$$E[\|X(k)\|^p] \leq C \gamma^{pk} \quad (10)$$

for every $k \geq 0$. It is well known [13] that

$$[(9) \text{ is } p\text{th mean stable}] \Leftrightarrow [\rho_p < 1]. \quad (11)$$

In this example, we let $N = 2$ and randomly choose two matrices

$$A_1 = \begin{bmatrix} -0.87 & -0.77 \\ 1.17 & -1.09 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0.14 & 0.40 \\ 0.89 & -0.73 \end{bmatrix}.$$

Using a MATLAB implementation [19] of the stochastic gradient descent algorithm, we perform the maximization over the set \mathfrak{D}_m and find

$$W_1 = \begin{bmatrix} -0.71 & -0.70 \\ 0.70 & -0.71 \end{bmatrix}, \quad W_2 = \begin{bmatrix} 0.85 & -0.53 \\ 0.53 & -0.85 \end{bmatrix}.$$

Since $\lambda_{\{W_1, W_2\}} = 1.07$ we obtain $\ell_2 \geq 1.07$. Thus, Theorem 3 implies that $\rho_1 > 1$ and hence the system in (9) is unstable, according to (11). We cannot prove that the system in (9) is unstable using the other lower bounds in the literature, since $\ell_Z = 0.93$ and $\ell_{[-1,1]} = 0.73$. We remark that the joint spectral radius appearing in ℓ_Z is evaluated with the JSR Toolbox [14]. Also, the maximum in $\ell_{[-1,1]}$ has been evaluated with extensive simulations on the weights.

We propose a further extension of our bounds based on product families of a set of matrices. Let us define \mathcal{M}^q to be the family of matrices consisting of all the N^q products of the matrices from \mathcal{M} having length q . Then, our extension can be stated as follows:

Theorem 9. *Let m and q be positive integers. Define $\ell_m^{(q)}(\mathcal{M}) = \ell_m(\mathcal{M}^q)^{1/q}$. Then, $\ell_m^{(q)} \leq \rho_1$. Furthermore, if q is a divisor of another positive integer q' , then $\ell_m^{(q)} \leq \ell_m^{(q')}$.*

Proof. Let us first recall the following identities [13]:

$$\rho_p(\mathcal{M}^q) = \rho_p(\mathcal{M})^q, \quad (12)$$

$$\rho_\infty(\mathcal{M}^q) = \rho_\infty(\mathcal{M})^q. \quad (13)$$

Then, using Theorem 3 and Equation (12), we can prove the first claim in the theorem as $\rho_1(\mathcal{M}) = \rho_1(\mathcal{M}^q)^{1/q} \geq \ell_m(\mathcal{M}^q)^{1/q} = \ell_m^{(q)}(\mathcal{M})$. To prove the second claim in the theorem, we let $q = 1$ and $q' = 2$ for simplicity. The proof

for general q and q' is similar to this particular case and hence is omitted. Let $\mathcal{W} = \{W_1, \dots, W_N\} \subset \mathbb{R}^{m \times m}$ be arbitrary and assume $\rho_\infty(\mathcal{W}) = 1$. Since $\rho(M)^2 = \rho(M^2)$ for a square matrix M , we can show that

$$\begin{aligned} \lambda_{\mathcal{W}}(\mathcal{M})^2 &= \rho \left(\left(\frac{1}{N} \sum_{i=1}^N W_i \otimes A_i \right)^2 \right) \\ &= \rho \left(\frac{1}{N^2} \sum_{i,j=1}^N (W_i W_j) \otimes (A_i A_j) \right) \\ &= \lambda_{\mathcal{W}^2}(\mathcal{M}^2). \end{aligned} \tag{14}$$

Also, since $\rho_\infty(\mathcal{W}^2) = \rho_\infty(\mathcal{W})^2 = 1$ by (13), it follows that $\lambda_{\mathcal{W}^2}(\mathcal{M}^2) \leq \ell(\mathcal{M}^2) = \ell^{(2)}(\mathcal{M})^2$. This inequality and (14) yield $\lambda_{\mathcal{W}} \leq \ell^{(2)}$. Taking the supremum with respect to \mathcal{W} in the left-hand side of this inequality proves $\ell_m^{(1)} \leq \ell_m^{(2)}$, as desired. \square

We close this section by giving a remark on complex weights.

Remark 10. In principle, one could obtain better lower bounds using complex weights in ℓ_m instead of the real weights W_1, \dots, W_N . However, we can show that it does not lead to an essential improvement. Precisely speaking, we here prove the following claim: if $\mathcal{W} = \{W_1, \dots, W_N\} \subset \mathbb{C}^{m \times m}$ satisfies $\rho_\infty(\mathcal{W}) = 1$, then

$$\lambda_{\mathcal{W}} \leq \ell_{2m}. \tag{15}$$

To prove this claim, for $W \in \mathbb{C}^{m \times m}$ we let

$$T_W = \begin{bmatrix} \operatorname{Re} W & -\operatorname{Im} W \\ \operatorname{Im} W & \operatorname{Re} W \end{bmatrix} \in \mathbb{R}^{(2m) \times (2m)},$$

where $\operatorname{Re} W$ and $\operatorname{Im} W$ denotes the real and imaginary parts of W , respectively. The multiplicative property $T_{WW'} = T_W T_{W'}$ and the identity $\|T_W\| = \|W\|$ yield

$$\rho_\infty(T_{W_1}, \dots, T_{W_N}) = 1. \tag{16}$$

Also, since $\rho(W) = \rho(T_W)$, we can show that

$$\begin{aligned}
\lambda_{\mathcal{W}}(\mathcal{M}) &= \rho\left(T_{N^{-1} \sum_{i=1}^N W_i \otimes A_i}\right) \\
&= \rho\left(\frac{1}{N} \begin{bmatrix} \sum_{i=1}^N (\operatorname{Re} W_i) \otimes A_i & \sum_{i=1}^N -(\operatorname{Im} W_i) \otimes A_i \\ \sum_{i=1}^N (\operatorname{Im} W_i) \otimes A_i & \sum_{i=1}^N (\operatorname{Re} W_i) \otimes A_i \end{bmatrix}\right) \\
&= \rho\left(\frac{1}{N} \sum_{i=1}^N T_{W_i} \otimes A_i\right) \\
&= \lambda_{\{T_{W_1}, \dots, T_{W_N}\}}(\mathcal{M}).
\end{aligned}$$

This equation and (16) prove Inequality (15).

3. The Markovian case

In this section, we extend the results presented in the last section to the Markovian case. Let $\sigma = \{\sigma(k)\}_{k=1}^\infty$ be a time-homogeneous Markov chain with a state space $\{1, \dots, N\}$ and a transition probability matrix $Q \in \mathbb{R}^{N \times N}$. We define the Markovian version of the p -radius as follows.

Definition 11. Let $X(\cdot; \mu)$ denote the trajectory of the Markov jump linear system

$$\Sigma : X(k+1) = A_{\sigma(k+1)} X(k), \quad X(0) = I, \quad \sigma(1) \sim \mu,$$

where μ is an arbitrary probability distribution on $\{1, \dots, N\}$. The L^p -norm Markovian joint spectral radius (Markovian p -radius for short) of the pair (\mathcal{M}, Q) is defined by

$$\rho_p(\mathcal{M}, Q) = \sup_{\mu} \limsup_{k \rightarrow \infty} \left(E[\|X(k; \mu)\|^p]^{1/(pk)} \right).$$

Remark 12. The logarithm of the Markovian p -radius corresponds to a quantity called Lyapunov exponent of the p th mean [20, p. 307].

We will later see that this definition coincides with the originally defined p -radius in the case where the matrices $A_{\sigma(k)}$ are independent and uniformly distributed at every time step. Moreover, as can be naturally expected, the Markovian p -radius has a close connection with the mean stability of Σ , which is defined as follows. We say that Σ is p th mean stable if there exist $C > 0$ and $\gamma \in [0, 1)$ such that (10) holds for all k and μ . The next proposition is an immediate consequence from the definition of the Markovian p -radius.

Proposition 13. Σ is p th mean stable if and only if $\rho_p(\mathcal{M}, Q) < 1$.

Proof. The necessity is obvious. Let us prove the sufficiency. Assume that $\rho_p(\mathcal{M}, Q) < 1$. Then there exists a $\gamma \in [0, 1)$ such that, for every $i \in \{1, \dots, N\}$, we have $\limsup_{k \rightarrow \infty} (E[\|X(k; \delta_i)\|^p]^{1/(pk)}) < \gamma$, where δ_i is the probability distribution on $\{1, \dots, N\}$ such that $\delta_i(\{i\}) = 1$. Then, for each i there exists a positive integer K_i such that, if $k > K_i$, then $E[\|X(k; \delta_i)\|^p] < \gamma^{pk}$. Therefore, for an arbitrary μ , if $k > \max(K_1, \dots, K_N)$ then $E[\|X(k; \mu)\|^p] = \sum_{i=1}^N \mu_i E[\|X(k; \delta_i)\|^p] \leq \sum_{i=1}^N \mu_i \gamma^{pk} = \gamma^{pk}$. This implies that Σ is p th mean stable. \square

The main result in this section is stated in the following theorem, which can be used to compute upper and lower bounds on the Markovian p -radius.

Theorem 14. 1. The sequence $\{h_k(\mathcal{M}, Q)\}_{k=1}^\infty$ defined by

$$h_k(\mathcal{M}, Q) = \left(\sum_{i_1, \dots, i_k=1}^N q_{i_1, i_2} \cdots q_{i_{k-1}, i_k} \|A_{i_k} \cdots A_{i_1}\|^p \right)^{1/(pk)}$$

is decreasing. Moreover $\rho_p(\mathcal{M}, Q) = \lim_{k \rightarrow \infty} h_k(\mathcal{M}, Q)$.

2. For $\bar{W} = \{W_{ij}\}_{1 \leq i, j \leq N} \subset \mathbb{R}^{m \times m}$ define

$$\mathcal{A}_{\bar{W}} = \begin{bmatrix} q_{11}W_{11} \otimes A_1 & \cdots & q_{N1}W_{N1} \otimes A_N \\ \vdots & \ddots & \vdots \\ q_{1N}W_{1N} \otimes A_1 & \cdots & q_{NN}W_{NN} \otimes A_N \end{bmatrix}.$$

If $\rho_\infty(\bar{W}) = 1$, then $\rho_1(\mathcal{M}, Q) \geq \rho(\mathcal{A}_{\bar{W}})$.

Let us observe some consequences of Theorem 14 before proving it. First, consider the special case when σ is a sequence of independent and uniformly distributed random variables on $\{1, \dots, N\}$. The corresponding transition probability matrix Q is the $N \times N$ matrix whose entries are all $1/N$. In this case, we have $h_k(\mathcal{M}, Q) = (\sum_{i_1, \dots, i_k=1}^N (1/N)^{k-1} \|A_{i_k} \cdots A_{i_1}\|^p)^{1/(pk)} = N^{1/kp} h_k(\mathcal{M})$. This equation implies that $h_k(\mathcal{M}, Q)$, the average of the norm of k -product $A_{i_k} \cdots A_{i_1}$, coincides with the other average $h_k(\mathcal{M})$ under the identification of Σ as a switched linear system having independent and identically distributed system parameters, except the factor $N^{1/kp}$. This factor,

roughly speaking, arises because $h_k(\mathcal{M}, Q)$ does not take the initial probability distribution of the switching signal σ into account. Then, by this equation and the first claim of Theorem 14, taking the limit as $k \rightarrow \infty$ shows that $\rho_p(\mathcal{M}, Q) = \rho_p(\mathcal{M})$. Hence Definition 11 indeed recovers the original p -radius.

Next, as a corollary of Theorem 14, we can recover a lower bound of the Markovian p -radius implicitly presented in [7].

Corollary 15 ([7]). *Define*

$$\ell_m(\mathcal{M}, Q) = \sup \left\{ \rho(\mathcal{A}_{\bar{\mathcal{W}}}) : \bar{\mathcal{W}} \in (\mathbb{R}^{m \times m})^{N^2}, \rho_\infty(\bar{\mathcal{W}}) = 1 \right\}.$$

Then, $\ell_1(\mathcal{M}, Q) \leq \rho_1(\mathcal{M}, Q)$.

Before we prove Theorem 14, we illustrate its use with an example.

Example 16. Consider the Markov jump linear system Σ with

$$A_1 = \begin{bmatrix} 0.77 & 0.80 \\ -0.60 & 0.87 \end{bmatrix}, \quad A_2 = \begin{bmatrix} -0.77 & 0.83 \\ -0.70 & -0.70 \end{bmatrix},$$

$$Q = \begin{bmatrix} 0.70 & 0.30 \\ 0.43 & 0.57 \end{bmatrix}.$$

A brute force search indicates that $\ell_1(\mathcal{M}, Q) = 0.844$ with the weights $W_{11} = 1$, $W_{12} = 1$, $W_{21} = -1$, and $W_{22} = 0.932$. On the other hand, using a MATLAB implementation [19] of the stochastic gradient descent algorithm used in Example 8, we locally maximize $\ell_2(\mathcal{M}, Q)$ over the set \mathfrak{D}_2 of matrix weights to find the set $\bar{\mathcal{W}}$ consisting of the following matrices:

$$W_{11} = \begin{bmatrix} -0.412 & -0.911 \\ 0.911 & -0.412 \end{bmatrix}, \quad W_{12} = \begin{bmatrix} 0.839 & -0.544 \\ 0.544 & -0.839 \end{bmatrix},$$

$$W_{21} = \begin{bmatrix} -0.204 & -0.979 \\ 0.979 & -0.204 \end{bmatrix}, \quad W_{22} = \begin{bmatrix} 0.937 & -0.349 \\ 0.349 & -0.937 \end{bmatrix}.$$

Since $\rho(\mathcal{A}_{\bar{\mathcal{W}}}) = 1.067$, we conclude $\rho_1(\mathcal{M}, Q) \geq 1.067$. This proves that the corresponding Markov jump linear system is unstable.

The rest of this section is devoted to the proof of Theorem 14. In our proof, we employ a reduction of the Markovian p -radius to the original p -radius, for which we can apply the results obtained in Section 2.

Proposition 17. *Let p be a positive integer. For each $1 \leq i \leq N$, let e_i denote the i th vector in the canonical basis of \mathbb{R}^N . Define the set of matrices $\bar{\mathcal{M}}_p = \{B_{ij}^{(p)}\}_{i,j \in \{1, \dots, N\}}$ as*

$$B_{ij}^{(p)} = N^{2/p} q_{ij}^{1/p} A_i \otimes (e_j e_i^\top).$$

Then,

$$\rho_p(\mathcal{M}, Q) = \rho_p(\bar{\mathcal{M}}_p). \quad (17)$$

Proof. We first claim that, for the proof of (17), it is sufficient to show that

$$[\rho_p(\mathcal{M}, Q) < 1] \Leftrightarrow [\rho_p(\bar{\mathcal{M}}_p) < 1], \quad (18)$$

due to the following reason. Suppose that (17) does not hold while (18) is true. Then, we have either $\rho_p(\mathcal{M}, Q) > \rho_p(\bar{\mathcal{M}}_p)$ or $\rho_p(\mathcal{M}, Q) < \rho_p(\bar{\mathcal{M}}_p)$. If the former inequality holds, then one can find a $c > 0$ such that the matrix family $c\mathcal{M} = \{cA_i\}_{i=1}^N$ satisfies $\rho_p(c\mathcal{M}, Q) > 1 > \rho_p(c\bar{\mathcal{M}}_p) = \rho_p(\bar{c}\bar{\mathcal{M}}_p)$. However, this cannot be true by (18). In a similar way, we can also show that the latter inequality, $\rho_p(\mathcal{M}, Q) < \rho_p(\bar{\mathcal{M}}_p)$, cannot hold. Therefore, (17) must be true.

Consequently, it suffices to prove (18). In order to prove the claim in (18), we introduce an alternative switched linear system with independent and identically distributed jumping parameters. Let $\{\theta(k)\}_{k=1}^\infty$ and $\{\phi(k)\}_{k=1}^\infty$ be independent random variables uniformly distributed on $\{1, \dots, N\}$. Define the switched linear system $\bar{\Sigma}_p$ by

$$\bar{\Sigma}_p : \bar{X}(k+1) = B_{\theta(k+1), \phi(k+1)}^{(p)} \bar{X}(k), \quad \bar{X}(0) = I.$$

From (11), we see that $\rho_p(\bar{\mathcal{M}}_p) < 1$ if and only if $\bar{\Sigma}_p$ is p th mean stable. Moreover, from Proposition 13, we know that $\rho_p(\mathcal{M}, Q) < 1$ if and only if Σ is p th mean stable. Therefore, to prove (18), we need to show that Σ is p th mean stable if and only if $\bar{\Sigma}_p$ is p th mean stable.

To prove the equivalence of stability, let us first compute $E[\|\bar{X}(k)\|^p]$. By the definition of $\bar{\Sigma}_p$ and Equation (2), we can compute $\bar{X}(k)$ as

$$\begin{aligned} \bar{X}(k) &= \prod_{i=1}^k \left(N^{2/p} q_{\theta(i), \phi(i)}^{1/p} A_{\theta(i)} \otimes (e_{\phi(i)} e_{\theta(i)}^\top) \right) \\ &= N^{2k/p} (q_{\theta(1), \phi(1)} \cdots q_{\theta(k), \phi(k)})^{1/p} (A_{\theta(k)} \cdots A_{\theta(1)}) \otimes J_k, \end{aligned}$$

where the symbol $\prod_{i=1}^k$ denotes the left product of matrices and also $J_k = e_{\phi(k)} e_{\theta(k)}^\top \cdots e_{\phi(1)} e_{\theta(1)}^\top$. Since the vectors e_1, \dots, e_N are orthonormal, $\|J_k\| = 1$ if

$$\phi(i) = \theta(i+1), \quad i = 1, \dots, k-1, \quad (19)$$

and otherwise $\|J_k\| = 0$. Therefore, if we denote by χ the characteristic function of the event (19), using (3) we obtain

$$\|\bar{X}(k)\|^p = \chi N^{2k} q_{\theta(1), \theta(2)} \cdots q_{\theta(k-1), \theta(k)} q_{\theta(k), \phi(k)} \|A_{\theta(k)} \cdots A_{\theta(1)}\|^p.$$

Since the event in (19) occurs with probability $1/N^{k-1}$, we obtain

$$\begin{aligned} & E[\|\bar{X}(k)\|^p] \\ &= \frac{N^{2k}}{N^{k-1}} E[q_{\theta(1), \theta(2)} \cdots q_{\theta(k-1), \theta(k)} q_{\theta(k), \phi(k)} \|A_{\theta(k)} \cdots A_{\theta(1)}\|^p] \\ &= N^{k+1} \sum_{i_1, \dots, i_k, j=1}^N \frac{1}{N^{k+1}} q_{i_1, i_2} \cdots q_{i_{k-1}, i_k} q_{i_k, j} \|A_{i_k} \cdots A_{i_1}\|^p \\ &= \sum_{i_1, \dots, i_k=1}^N q_{i_1, i_2} \cdots q_{i_{k-1}, i_k} \|A_{i_k} \cdots A_{i_1}\|^p \\ &= \sum_{i_1=1}^N E[\|X(k; \delta_{i_1})\|^p], \end{aligned} \quad (20)$$

where we used $\sum_{j=1}^N q_{i_k, j} = 1$ to show the third equality.

Now, assume that Σ is p th mean stable. Then, by (20) we have $E[\|\bar{X}(k)\|^p] \leq \sum_{i_1=1}^N C \gamma^{pk} = CN \gamma^{pk}$ for some $C > 0$ and $\gamma \in [0, 1)$, and hence $\bar{\Sigma}_p$ is p th mean stable. On the other hand, assume that $\bar{\Sigma}_p$ is p th mean stable. Then, there exist $C > 0$ and $\gamma \in [0, 1)$ such that $E[\|\bar{X}(k)\|^p] \leq C \gamma^{pk}$. Therefore, (20) shows that $E[\|X(k; \delta_{i_1})\|^p] \leq C \gamma^{pk}$ for every $i_1 \in \{1, \dots, N\}$. Then, in the same way as the proof of Proposition 13, we can conclude the p th mean stability of Σ . This completes the proof of the theorem. \square

Remark 18. From Proposition 17, we can regard the matrix $B_{ij}^{(p)}$ as an L^p -averaged version of the Ω -lift introduced in [11], which is used to generalize the so-called Berger-Wang formula to a Markovian version of the joint spectral radius. Also, we remark that considering the auxiliary switched linear

system $\bar{\Sigma}_p$ with an extended state space is similar to considering an extended state-variable consisting of the original state variable and the underlying Markov chain, which is frequently employed for studying (semi-)Markov jump linear systems [10] ([21]).

Finally, let us prove Theorem 14.

Proof of Theorem 14. From Equation (20), we see that

$$h_k(\bar{\mathcal{M}}_p) = E[\|\bar{X}(k)\|^p]^{1/(pk)} = h_k(\mathcal{M}, Q).$$

Therefore, by Proposition 1, the sequence $\{h_k(\mathcal{M}, Q)\}_{k=1}^\infty$ is decreasing. Furthermore, it converges to $\lim_{k \rightarrow \infty} h_k(\bar{\mathcal{M}}_p) = \rho_p(\bar{\mathcal{M}}_p) = \rho_p(\mathcal{M}, Q)$ by (17). Thus, the first claim is proved. Let us then prove the second claim. Assume that $\bar{\mathcal{W}} = \{W_{ij}\}_{1 \leq i, j \leq N} \subset \mathbb{R}^{m \times m}$ satisfies $\rho_\infty(\bar{\mathcal{W}}) = 1$. By (17), it is enough to show that $\rho_1(\bar{\mathcal{M}}_1) \geq \rho(\mathcal{A}_{\bar{\mathcal{W}}})$. From Theorem 3 we obtain

$$\begin{aligned} \rho_1(\bar{\mathcal{M}}_1) &\geq \rho \left(\frac{1}{N^2} \sum_{i,j=1}^N W_{ij} \otimes B_{ij}^{(1)} \right) \\ &= \rho \left(\sum_{i,j=1}^N q_{ij} W_{ij} \otimes A_i \otimes (e_j e_i^\top) \right). \end{aligned} \tag{21}$$

Here we recall that there exists [12] an invertible matrix T satisfying $C \otimes D = T^{-1}(D \otimes C)T$ for all $C \in \mathbb{R}^{N \times N}$ and $D \in \mathbb{R}^{(nm) \times (nm)}$. Consequently, there exists an invertible matrix T such that $W_{ij} \otimes A_i \otimes (e_j e_i^\top) = T^{-1}((e_j e_i^\top) \otimes W_{ij} \otimes A_i)T$. Therefore, the matrix appearing in the last term of (21) is similar to $\sum_{i,j=1}^N q_{ij} (e_j e_i^\top) \otimes W_{ij} \otimes A_i$, which in fact equals $\mathcal{A}_{\bar{\mathcal{W}}}$. Therefore $\rho_1(\bar{\mathcal{M}}_1) \geq \rho(\mathcal{A}_{\bar{\mathcal{W}}})$, as desired. \square

4. Conclusion

This paper proposed novel lower bounds on the p -radius of a finite set of matrices. The obtained lower bound is given by the spectral radius of an average of the given matrices weighted via Kronecker products. We showed that the proposed lower bounds theoretically extend and also practically improve the existing lower bounds. We have also shown the extension of the p -radius and its lower bounds to the Markovian case.

Acknowledgment

This work was supported in part by the NSF under grants CNS-1302222 and IIS-1447470. R.J. is an FNRS Research Associate. His work is supported by the Communauté française de Belgique (ARC), and by the Belgian state (PAI). Parts of this work were carried out while the first author was visiting the Institute of ICTEAM (Information and Communication Technologies, Electronics and Applied Mathematics) at Université catholique de Louvain. He would like to acknowledge the generous hospitality of the institute.

References

- [1] K.-S. Lau, J. Wang, Characterization of L^p -solutions for the two-scale dilation equations, *SIAM Journal on Mathematical Analysis* 26 (4) (1995) 1018–1046.
- [2] V. Y. Protasov, The generalized joint spectral radius. A geometric approach, *Izvestiya: Mathematics* 61 (5) (1997) 995–1030.
- [3] R.-Q. Jia, Subdivision schemes in L_p spaces, *Advances in Computational Mathematics* 3 (4) (1995) 309–341.
- [4] Y. Wang, Two-scale dilation equations and the mean spectral radius, *Random and Computational Dynamics* 4 (1996) 49–72.
- [5] M. Ogura, C. F. Martin, Generalized joint spectral radius and stability of switching systems, *Linear Algebra and its Applications* 439 (8) (2013) 2222–2239.
- [6] R. M. Jungers, V. Y. Protasov, Weak stability of switching dynamical systems and fast computation of the p -radius of matrices, in: 49th IEEE Conference on Decision and Control, IEEE, 2010, pp. 7328–7333.
- [7] J. Barthélemy, M. Marx, Monetary policy switching and indeterminacy, Working paper.
- [8] R. M. Jungers, V. Y. Protasov, Fast methods for computing the p -radius of matrices, *SIAM Journal on Scientific Computing* 33 (3) (2011) 1246–1266.

- [9] D.-X. Zhou, The p -norm joint spectral radius for even integers, *Methods and Applications of Analysis* 5 (1) (1998) 39–54.
- [10] O. Costa, M. Fragoso, R. Marques, *Discrete-Time Markov Jump Linear Systems, Probability and Its Applications*, Springer-Verlag, London, 2005.
- [11] V. Kozyakin, The Berger–Wang formula for the Markovian joint spectral radius, *Linear Algebra and its Applications* 448 (13) (2014) 315–328.
- [12] J. Brewer, Kronecker products and matrix calculus in system theory, *IEEE Transactions on Circuits and Systems* 25 (9) (1978) 772–781.
- [13] R. M. Jungers, *The Joint Spectral Radius*, Vol. 385 of *Lecture Notes in Control and Information Sciences*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [14] G. Vankeerberghen, J. Hendrickx, R. M. Jungers, JSR: a toolbox to compute the joint spectral radius, in: *Proceedings of the 17th international conference on Hybrid systems: computation and control - HSCC '14*, ACM Press, New York, New York, USA, 2014, pp. 151–156.
- [15] M. L. Overton, R. S. Womersley, On minimizing the special radius of a nonsymmetric matrix function: optimality conditions and duality theory, *SIAM Journal on Matrix Analysis and Applications* 9 (4) (1988) 473–498.
- [16] W. Hürlimann, Generalized Helmert-Ledermann orthogonal matrices and rom simulation, *Linear Algebra and its Applications* 439 (7) (2013) 1716–1729.
- [17] J. V. Burke, A. S. Lewis, M. L. Overton, A robust gradient sampling algorithm for nonsmooth, nonconvex optimization, *SIAM Journal on Optimization* 15 (3) (2005) 751–779.
- [18] F. Kozin, A survey of stability of stochastic systems, *Automatica* 5 (1) (1969) 95–112.
- [19] J. Burke, A. Lewis, M. Overton, *Hybrid Algorithm for Non-Smooth Optimization*, a MATLAB package for non-smooth non-convex optimization (2006).

- [20] Y. Fang, K. A. Loparo, X. Feng, Stability of discrete time jump linear systems, *Journal of Mathematical Systems, Estimation, and Control* 5 (3) (1995) 275–321.
- [21] M. Ogura, C. F. Martin, Stability analysis of positive semi-Markovian jump linear systems with state resets, *SIAM Journal on Control and Optimization* 52 (3) (2014) 1809–1831.